Beliakov, Gleb and James, Simon 2013, Stability of weighted penalty-based aggregation functions, Fuzzy sets and systems, vol. 226, pp. 1-18.

Stability of weighted penalty-based aggregation functions

Gleb Beliakov[*], Simon James[†]

January 5, 2013

Abstract

In many practical applications, the need arises to aggregate data of varying dimension. Following from the self-identity property, some recent studies have looked at the stability of aggregation operators in terms of their behavior as the dimensionality is increased from $n-1$ to n. We use the penalty-based representation of aggregation functions in order to investigate the conditions for weighting vectors associated with some important weighted families, extending on the results already established for quasi-arithmetic means. In particular, we obtain results for quasi-medians and functions that involve a reordering of the inputs such as the OWA and order statistics.

Key words: Aggregation functions, means, medians, quasi-arithmetic means, OWA, stability, weighting triangles.

1 Introduction

A problem that arises in decision making and information fusion is how to deal with data of varying dimension. If comparing two items based on multiple criteria or a group of experts' opinions, it may be that some evaluations are missing. Similarly, when fusing the readings of sensors, it could be that not all of the readings are available at all times and we need a global evaluation based on some subset. Thus we are

[*]email - gleb@deakin.edu.au

[†]email - sjames@deakin.edu.au , Ph +61 3 9251 7481 (corresponding author)

looking for families of aggregation functions that produce consistent outputs regardless of input cardinality.

Some aggregation functions, such as t-norms, t-conorms and uninorms, are associative, and therefore have a natural way of defining n-variable instances. Quasi-arithmetic means (with equal weights) is another example where the whole family of functions is defined consistently. In contrast, defining weighted means and ordered weighted averaging (OWA) operators consistently does represent a significant challenge.

How to define families of weighted aggregation functions has been approached in [6,8–10] with the construction of weighting triangles and the notion of *extended* aggregation functions. Methods include defining recursive sequences of weighting vectors and the use of quantifiers (which have been especially important for OWA families [21, 22]).

The use of weighting triangles allows a kind of mathematical consistency between members of a family of aggregation functions defined for varying dimension, however some recent studies have also analyzed such families from the viewpoint of stability [16, 19]. Whilst stability and robustness of aggregation is usually thought of in terms of concepts like Lipschitz continuity [2, 11] (a small increase to one of the inputs should not result in a drastic increase to the output), it also makes sense that the inclusion of an additional input should not drastically alter the aggregated value if it is representative of the rest. From Yager's self-identity property [24] the authors of [16, 19] consider the stability of various classes of aggregation functions. The idea is that given a set of inputs, if we add the aggregated value of these inputs as a new input, the overall output should not change. A function is considered to be stable if the new input can be aggregated both from the right and the left (i.e., as either the last or first argument respectively).

This behavior was referred to as $F-$ insensitivity by Gągolewsk and Grzegorzewski in the context of extended aggregation operators satisfying a property they refer to as *arity-monotonicity* [14,15]. Arity-monotone functions are non-decreasing with the addition of a new input, which for the producer assessment problem ensures that increased productivity does not result in a lower overall evaluation. In particular, they looked at conditions on the weights for symmetric functions that involve a reordering of the inputs such as OWA operators and the ordered weighted maximum.

In this work, we adopt the definitions from [16, 19], and firstly, extend their results to all quasi-arithmetic means, quasi-medians and other averaging functions that can be defined as penalty-based functions [7]. We then consider weighted aggregation functions in detail.

For weighted functions, the i-th input is usually representative of the source, particular criterion, expert etc. When we increase the dimension, the additional inputs would usually be added in the last position, so we will focus on the notion of R-strict stability. Results for j-th position stability, however can be easily obtained as a corollary with a re-indexing of the weights. We draw upon the notion of penalty-based aggregation functions to explore the stability of different aggregation function families. By using the penalty-based expressions, the results that we establish can be applied broadly to a number of important cases.

We also investigate penalties expressed with a reordering of the inputs. As well as the stability of weighted aggregation functions being important conceptually for robustness of the aggregation process, a recent example of the need for such conventions is in defining the generalized Bonferroni mean [3] which uses two means of n and $n-1$ arguments respectively in its construction.

The article will be set out as follows. In Section 2 we will provide the preliminary notions required for the rest of the paper, including aggregation functions, penalties, and the definitions of strict-stability [16,19]. We also present a useful corollary of the propositions in [16,19] regarding the relationship between weighting vectors. In Section 3, we propose the necessary conditions on weighted penalty functions that lead to R-strictly stable aggregation functions. In Section 4 we turn to aggregation functions that are calculated with a reordering of the inputs, before summarizing our findings in the final section.

2 Preliminaries

We will approach stability from the viewpoint of constructing penalty-based aggregation functions. We first give an overview of aggregation functions, their families and some properties including strict stability, and then show how these notions relate to their penalty-based representations.

2.1 Aggregation functions

Aggregation functions take multiple arguments and combine them into a single value which is seen to be representative. Their properties, construction methods and applications have been investigated in the recent monographs [4, 17, 20]. We will consider aggregation functions defined over the unit interval.

Definition 1 *An aggregation function* $f : [0,1]^n \to [0,1]$ *is a function non-decreasing in each argument and satisfying* $f(0,\ldots,0) = 0$ *and* $f(1,\ldots,1) = 1$.

Depending on the application, further properties and behavior are often desired. In particular, we are interested in averaging aggregation functions, which can be defined in terms of their minimum and maximum arguments.

Definition 2 *An aggregation function* f *is considered to be* averaging *when*

$$\min(\mathbf{x}) \leq f(\mathbf{x}) \leq \max(\mathbf{x}),$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$.

Due to the monotonicity of aggregation functions, averaging behavior is equivalent to idempotency, i.e. $f(t, t, ..., t) = t$.

Typical examples of averaging aggregation functions include the arithmetic mean (also referred to as the statistical *average*),

$$AM(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

and the median,

$$Med(\mathbf{x}) = \begin{cases} x_{(k)} , & n = 2k-1, \\ 0.5(x_{(k)} + x_{(k+1)}) , & n = 2k, \end{cases}$$

where $x_{(k)}$ denotes the k-th input when the inputs are arranged into non-decreasing order.

Note that these functions are defined for all n, although the parameters will change as the dimension varies, e.g. the weight applied to each input when $n = 4$ will be $1/4$ for the arithmetic mean, while for $n = 50$ it will be $1/50$.

An important generalized family of averaging functions are the weighted quasi-arithmetic means.

Definition 3 *For a strictly monotone continuous generating function* $\phi : [0,1] \to [-\infty, \infty]$ *and weighting vector* $\mathbf{w}$*, the weighted quasi-arithmetic mean is given by,*

$$QAM_{\mathbf{w}}(\mathbf{x}) = \phi^{-1}\left(\sum_{i=1}^{n} w_i \phi(x_i)\right). \quad (1)$$

Special cases include:

$\phi(t) = t$, the weighted arithmetic mean, $WAM(\mathbf{x}) = \sum_{i=1}^{n} w_i x_i$;

$\phi(t) = t^q$, the weighted power mean, $PM_q(\mathbf{x}) = \left(\sum_{i=1}^{n} w_i x_i^q\right)^{1/q}$;

$\phi(t) = -\ln t$, the weighted geometric mean, $G(\mathbf{x}) = \prod_{i=1}^{n} x_i^{w_i}$.

It is usually required that the weights are non-negative with $\sum_{i=1}^{n} w_i = 1$. For equal weights, we recover the symmetric cases corresponding with each generator.

The weight w_i often is indicative of the importance of the input x_i. For instance, in multi-criteria decision making, it may be that criterion 1 is more important than the others, so we ensure that w_1 is the largest weight.

A weighted extension of the median also exists.

Definition 4 *Given a weighting vector* $\mathbf{w}$*, we denote the corresponding vector* $\mathbf{u}$ *by rearranging the components of* $\mathbf{w}$ *according to a non-increasing permutation of the input vector* $\mathbf{x}$*, i.e.* $u_k = w_i$ *if* $x_i = x_{(k)}$ *is the* k*-th largest input . The lower weighted median is then given by*

$$Med_{\mathbf{w}}(\mathbf{x}) = x_{(k)}$$

where k *is the index obtained from the condition,*

$$\sum_{j=1}^{k-1} u_j < \frac{1}{2} \text{ and } \sum_{j=1}^{k} u_j \geq \frac{1}{2}.$$

For the upper weighted median, we exchange the inequalities $<, \geq$ *with* $\leq, >$ *respectively.*

OWA functions and their generalizations are also well known examples of averaging functions. Rather than allocate a weight to the i-th input, the weight assigned depends on the relative order of the inputs.

Definition 5 *Given a weighting vector* $\mathbf{w}$*, the OWA function is*

$$OWA_{\mathbf{w}}(\mathbf{x}) = \sum_{i=1}^{n} w_i x_{(i)},$$

where the (i) *notation denotes the components of* $\mathbf{x}$ *being arranged in non-increasing order* $x_{(1)} \geq x_{(2)} \geq \ldots \geq x_{(n)}$.

Special cases of the OWA operator, depending on the weighting vector $\mathbf{w}$, include the arithmetic mean where all the weights are equal, i.e. all $w_i = \frac{1}{n}$, the maximum for $\mathbf{w} = (1, 0, ..., 0)$, minimum for $\mathbf{w} = (0, ..., 0, 1)$ and the median, with $w_j = 1$ if $n = 2j - 1$ (n odd) or $w_j = w_{j+1} = 0.5$ if $n = 2j$ (n even) , and $w_i = 0$ otherwise. The OWA can also be used to model k-order statistics, with $w_k = 1$, and $w_i = 0$ for all $i \neq k$.

An important generalization of the OWA is the induced OWA [23], where the ordering of the input vector is determined by an auxiliary variable $\mathbf{z}$.

Definition 6 *Given a weighting vector* $\mathbf{w}$ *and an inducing variable* $\mathbf{z}$*, the Induced Ordered Weighted Averaging (IOWA) function is*

$$IOWA_{\mathbf{w}}(\langle x_1, z_1 \rangle, \ldots, \langle x_n, z_n \rangle) = \sum_{i=1}^{n} w_i x_{\sigma(i)}, \tag{2}$$

where the $\sigma(i)$ *notation denotes the inputs* $\langle x_i, z_i \rangle$ *reordered such that* $z_{\sigma(1)} \geq z_{\sigma(2)} \geq \ldots \geq z_{\sigma(n)}$ *and the convention that if* q *of the* z_i *are tied,*
i.e. $z_{\sigma(i)} = z_{\sigma(i+1)} = \ldots = z_{\sigma(i+q-1)}$,

$$x_{\sigma(i)} = \ldots = x_{\sigma(i+q-1)} = \frac{1}{q} \sum_{j=i}^{i+q-1} x_{\sigma(j)}.$$

For the preceding weighted functions, $\mathbf{w}$ needs to be specified for each n. This may be fixed for many applications, however if we consider a family of functions with varying dimension, the concept of a weighting triangle becomes useful for referring to the sequence of weighting vectors as n is increased.

Definition 7 *[10] A weighting triangle* W *is a sequence of weighting vectors,* $\mathbf{w}^n$*,* $n = 2, 3, \ldots$ *such that* $\sum_{i=1}^{n} w_i^n = 1$ *for each* n*. It can be represented by*

$$\begin{array}{cccc} & & 1 & & \\ & w_1^2 & & w_2^2 & \\ & w_1^3 & w_2^3 & w_3^3 & \\ w_1^4 & w_2^4 & w_3^4 & w_4^4 \\ & & \dots & & \end{array}$$

As we are interested in the stability of weighted functions as n varies, we will often make reference to the relationship between $\mathbf{w}^{n-1}$ and $\mathbf{w}^n$. We now turn to this notion of stability.

2.2 Strictly stable families

Following from Yager's self-identity property [24], Rojas et al. propose the following conditions for stability of a family of aggregation functions [19].

Definition 8 *Let $\{A_n : [0,1]^n \to [0,1], n \in N\}$ be a family of aggregation functions. Then it is said that:*

1. *$\{A_n\}_n$ is R-strictly stable if*

$$A_n(x_1, \dots, x_{n-1}, A_{n-1}(x_1, \dots, x_{n-1})) = A_{n-1}(x_1, \dots, x_{n-1}),$$

2. *$\{A_n\}_n$ is L-strictly stable if*

$$A_n(A_{n-1}(x_1, \dots, x_{n-1}), x_1, \dots, x_{n-1}) = A_{n-1}(x_1, \dots, x_{n-1}),$$

3. *$\{A_n\}_n$ is LR-strictly stable if both properties hold simultaneously.*

Rojas et al. established that geometric means and arithmetic means with respect to a weighting vector with equal weights, the maximum, minimum, and median are LR-strictly stable, while the weighted counterparts of these means and the OWA, in general, are unstable.

We emphasize that as is clear from the case of the OWA, the symmetry of an aggregation function does not imply strict-stability. The following example of T-S functions helps illustrate this.

Example 1 *Consider the T-S functions first defined in [25] and later studied in [18], which take the average of a t-norm (T) and t-conorm (S). The linear convex T-S function is given by,*

$$L_{\gamma,T,S}(\mathbf{x}) = (1 - \gamma) \cdot T(\mathbf{x}) + \gamma \cdot S(\mathbf{x}),$$

with $\gamma \in]0, 1[$.

If T is the minimum and S is the maximum, L will be strictly stable. Assume the inputs are in non-increasing order, we let $A_{n-1}(x_1, x_2, \ldots, x_{n-1}) = y^*$, with

$$\begin{aligned} y^* &= (1-\gamma) \cdot \min(x_1, x_2, \ldots, x_{n-1}) + \gamma \cdot \max(x_1, x_2, \ldots, x_{n-1}) \\ &= (1-\gamma)x_{n-1} + \gamma \cdot x_1. \end{aligned}$$

Since $x_{n-1} \leq y^* \leq x_1$, if we include y^* in the aggregation, we have

$$\begin{aligned} &A_n(x_1, x_2, \ldots, x_{n-1}, y^*) \\ &= (1-\gamma) \cdot \min(x_1, x_2, \ldots, x_{n-1}, y^*) + \gamma \cdot \max(x_1, x_2, \ldots, x_{n-1}, y^*) \\ &= (1-\gamma)x_{n-1} + \gamma \cdot x_1 \\ &= y^* = A_{n-1}(x_1, x_2, \ldots, x_{n-1}) \end{aligned}$$

as required.

However, it will not usually be the case that

$$T(x_1, x_2, \ldots, x_{n-1}, y^*) = T(x_1, x_2, \ldots, x_{n-1})$$

and

$$S(x_1, x_2, \ldots, x_{n-1}, y^*) = S(x_1, x_2, \ldots, x_{n-1}),$$

so other T-S functions will be unstable in general.

The strict stability conditions can be considered recursively for the definitions of aggregation functions of n and $n-1$ dimensions. In [13] some rules for defining such functions using a sequence of 2-variate weighted means were considered from the viewpoint of consistency and computability. The following proposition was established for weighted geometric, arithmetic and harmonic means in [16], and previously for weighting triangles associated with weighted quasi-arithmetic means in [8].

Proposition 1 *Let W be a weighting triangle according to Definition 7. The family of weighted means defined by these weights is R-strictly stable if and only if for all $n > 2$ the following holds:*

$$w_i^n = (1 - w_n^n) \cdot (w_i^{n-1}), i = 1, \ldots, n-1,$$

and L-strictly stable if and only if

$$w_i^n = (1 - w_1^n) \cdot (w_{i-1}^{n-1}), i = 2, \ldots, n.$$

An R-strictly stable weighting triangle is hence defined completely from the sequence $w_2^2, w_3^3, \ldots, w_n^n$. We can also consider the notion of j-th position stability, e.g. the case where additional inputs are always aggregated in the 2nd position.

Corollary 1 *A family of weighted means defined with respect to a weighting triangle W is j-th position-strictly stable if and only if for all $n > j$ it holds:*

$$w_i^n = \begin{cases} (1 - w_j^n) \cdot w_i^{n-1}, i < j, \\ (1 - w_j^n) \cdot w_{i-1}^{n-1}, i > j. \end{cases}$$

Proof. For $i < j$, corresponding weights for weighting vectors of n and $n-1$ dimensions will have the same index i. The new input is then inserted in the j-th position, shifting the index for all $i > j$, so for these inputs w_i^n will correspond with w_{i-1}^{n-1}. □

We also might be interested in conditions for weighting vectors to be both L- and R-strictly table. We obtain the following corollary from Proposition 1.

Corollary 2 *For weighted means, a weighting triangle is LR-strictly stable if and only if there exists a $\lambda \geq 0$ such that $w_i^n = \lambda w_{i-1}^n$ for $i = 2, 3, \ldots, n$.*

Proof. We express the 2-dimensional weighting vector in terms of λ, with $\mathbf{w}^2 = (w, \lambda w)$. For L-strict stability we require the ratio between the 2nd and 3rd input to be the same, since they both are determined from $\mathbf{w}^2$ by multiplying by $(1 - w_1^3)$, so

$$w_2^3 : w_3^3 = w : \lambda w$$

while similarly for R-strict stability we require

$$w_1^3 : w_2^3 = w : \lambda w.$$

For these to hold simultaneously, we recursively ensure that the ratio $w : \lambda w$ holds and it follows that the 3-dimensional weighting vector must have the ratio

$$w_1^3 : w_2^3 : w_3^3 = w : \lambda w : \lambda^2 w.$$

Since we require $w(1 + \lambda + \lambda^2) = 1$, the value of λ follows from the solution to

$$\lambda = -1 \pm \sqrt{1 - 4\left(1 - \frac{1}{w}\right)}$$

which has a unique feasible solution for all $0 < w \leq 1$.

As n increases to 4, we will require $w : \lambda w : \lambda^2 w : \lambda^3 w$ and so on.
□

This means we can determine all $\mathbf{w}^n$ of an LR-strictly stable weighting triangle from the ratio between the weights for $\mathbf{w}^2$. The value of $\lambda = 1$ leads to the weighting triangle where each $\mathbf{w}^n$ has equal weights. If $\lambda = 0$ we have the weighting triangle with $w_1^n = 1, w_i^n = 0$ otherwise, for all n, while the limiting case of $\lambda = \infty$ will have $w_n^n = 1, w_i^n = 0$ otherwise. If we have $\lambda = 2$, for example, the weighting triangle will be

$$\begin{array}{cccc} & & 1 & & \\ & \frac{1}{3} & & \frac{2}{3} & \\ \frac{1}{7} & & \frac{2}{7} & & \frac{4}{7} \\ \frac{1}{15} & \frac{2}{15} & & \frac{4}{15} & \frac{8}{15} \\ & & \cdots & & \end{array}$$

We will consider penalty-based aggregation functions to determine properties on the weights for weighted functions.

2.3 Penalty-based aggregation functions

The use of penalties to define important families of aggregation functions has been studied recently in [5, 7]. It was already well known to Laplace that arithmetic means and medians minimize respectively the squared and absolute differences between inputs and output [20], so it is useful to approach the construction of aggregation functions from this angle when we are interested in an aggregated value that is representative of the arguments.

Definition 9 *A penalty function* $P : [0,1]^{n+1} \to \bar{\Re}_+ = [0, \infty]$ *satisfies:*

i) $P(\mathbf{x}, y) \geq 0$ *for all* $\mathbf{x}, y$*;*

ii) $P(\mathbf{x}, y) = 0$ *if* $x_i = y \ \forall i$*;*

iii) For every fixed $\mathbf{x}$*, the set of minimizers of* $P(\mathbf{x}, y)$ *is either a singleton or an interval.*

The penalty based function is then given by

$$f(\mathbf{x}) = \arg\min_y P(\mathbf{x}, y),$$

if y is the unique minimizer, and $y = \frac{a+b}{2}$ if the set of minimizers is the interval (a, b) (open or closed).

Condition iii) can be satisfied by ensuring that $P(\mathbf{x}, y)$ is quasiconvex in y for any fixed $\mathbf{x}$.

A special class of penalty functions referred to as faithful penalty functions was investigated by Calvo et al. in [12]. Let P be given by

$$P(\mathbf{x}, y) = \sum_{i=1}^{n} w_i p(x_i, y), \tag{3}$$

where $p : [0, 1]^2 \to \Re_+$ is a dissimilarity function (or penalty) with the properties

1) $p(x, y) = 0$ if and only if $x = y$, and

2) $p(x, z) \geq p(y, z)$ whenever $x \geq y \geq z$ or $x \leq y \leq z$,

and $\mathbf{w} = (w_1, \ldots, w_n)$ is a weighting vector.

Remark 1 *In the case of faithful penalty functions, any positive scalar multiple of $\mathbf{w}$ will have the same minimizer and hence corresponds to an averaging aggregation function with weights given by $\frac{w_i}{\sum_{i=1}^{n} w_i}$. Where we have unequal weights, we will assume that $\sum_{i=1}^{n} w_i = 1$ so that each w_i will (usually) correspond with the associated weight in the resulting aggregation function.*

We list some of the special cases that will be of interest to us:

1. Let $p(x_i, y) = (x_i - y)^2$. The corresponding faithful penalty-based aggregation function is a weighted arithmetic mean;
2. Let $p(x_i, y) = |x_i - y|$. The corresponding faithful penalty-based aggregation function is a weighted median;
3. Let $P(\mathbf{x}, y) = \sum_{i=1}^{n} w_i p(x_{(i)}, y)$, where $x_{(i)}$ is the i-th largest component of $\mathbf{x}$. We obtain the ordered weighted counterparts of the means in the previous examples, namely the OWA and weighted medians. Where the order is induced by some auxiliary variable $\mathbf{z}$, we obtain the IOWA;
4. Let $c \geq 0$ and

$$p(x, y) = \begin{cases} x - y, & if\ x \leq y, \\ c(y - x), & if\ x > y. \end{cases} \tag{4}$$

   We obtain the α-quantile operator, with $\alpha = c/(1+c)$. To obtain the k-th order statistic, we take $c = \frac{i-1/2}{n-1+1/2}$;

5. Let $p(x,y) = (g(x) - g(y))^2$. The corresponding faithful penalty-based aggregation function is a weighted quasi-arithmetic mean with the generator g. The generalized OWA, generalized Choquet and quasi-medians can also be obtained by making the analogous substitutions ;
6. Let $p(x,y) = |g(x) - g(y)|$. The corresponding faithful penalty-based aggregation function is a weighted quasi-median with the generator g and defined as $f(\mathbf{x}) = g^{-1}(Med_{\mathbf{w}}(g(\mathbf{x})))$.

3 R-Strict stability

From Proposition 1 it follows that a family of weighted means cannot be consistently defined for all $n \geq 2$ such that it is both L- and R-strictly stable unless every pair of sequential weights satisfies $w_i = \lambda w_{i-1}$. On the other hand, the usual interpretation of weighting vectors is that the weight w_i reflects the importance of the input x_i, so it may not always make sense in applications to shift the indices of the inputs the way we do when A_{n-1} is aggregated in the first or j-th position. We may wish to insert a new input that is naturally ordered before some of those already included, however there is no reason why this value couldn't be aggregated in the n-th position as long as this is kept in mind when it comes to interpretation.

We will hence restrict the following considerations to the notion of R-strict stability, which is equivalent to the self-identity property, however results for any position j could be obtained with a simple re-indexing of the weights[1]. We will do this by means of penalty-based aggregation operators of the form given in Eq. (3). Expressing the functions in this way allows us to generalize the results for a number of important aggregation families, including weighted quasi-arithmetic means and weighted quasi-medians.

We use the notation $\mathbf{x}$ and $\mathbf{x}_{i \neq n}$ to denote the respective input vectors $(x_1, x_2, \ldots, x_{n-1}, x_n)$ and $(x_1, x_2, \ldots, x_{n-1})$. For aggregation functions expressed in terms of their penalties, we have the following propositions. Proposition 2 provides sufficient conditions on the weighting vectors for the aggregation family to be R-strictly stable, while Proposition 3 gives the necessary and sufficient conditions in the case that the penalty function is differentiable.

[1] We will also consider functions whose calculation can involve a reordering of the inputs in Section 4.

Proposition 2 *Given a family $\{A_n\}_n$ of faithful penalty-based aggregation operators with (Eq. (3))*

$$P(\mathbf{x}, y) = \sum_{i=1}^{n} w_i p(x_i, y),$$

if the weighting vectors $\mathbf{w}^n$ and $\mathbf{w}^{n-1}$ associated with the penalty expressions for each $A_n(\mathbf{x})$ and $A_{n-1}(\mathbf{x}_{i\neq n})$ satisfy:

$$w_i^n = \lambda_n w_i^{n-1}, i = 1, \ldots, n-1,$$

each $\lambda_n \geq 0$ a constant[2]*, then the family is R-strictly stable.*

Proof. For $A_n(\mathbf{x})$ and $A_{n-1}(\mathbf{x}_{i\neq n})$ with respect to the weighting vectors $\mathbf{w}^n$ and $\mathbf{w}^{n-1}$ respectively, we have the following penalty-based expressions.

$$\begin{aligned} A_n(\mathbf{x}) &= \arg\min_y \sum_{i=1}^{n} w_i^n p(x_i, y), \quad (5) \\ A_{n-1}(\mathbf{x}_{i\neq n}) &= \arg\min_y \sum_{i=1}^{n-1} w_i^{n-1} p(x_i, y). \quad (6) \end{aligned}$$

R-strict stability requires $A_n(x_1, \ldots, x_{n-1}, A_{n-1}(\mathbf{x}_{i\neq n})) = A_{n-1}(\mathbf{x}_{i\neq n})$. If we denote the minimizer in Eq. (6) by $y^* = A_{n-1}(\mathbf{x}_{i\neq n})$, this requirement can be stated in terms of the penalty-based expression of $A_n(\mathbf{x})$ as:

$$\arg\min_y \left(\sum_{i=1}^{n-1} w_i^n p(x_i, y) + w_n^n p(y^*, y) \right) = y^*. \quad (7)$$

The minimizer on the left hand side must be $y = y^*$ from which it follows that $p(y^*, y) = 0$. We hence discard the penalty associated with w_n^n and re-write Eq. (7) using the penalty expression of $y^* = A_{n-1}(\mathbf{x}_{i\neq n})$ on the right hand side. This gives

$$\arg\min_y \sum_{i=1}^{n-1} w_i^n p(x_i, y) = \arg\min_y \sum_{i=1}^{n-1} w_i^{n-1} p(x_i, y). \quad (8)$$

It is clear that if $w_i^n = \lambda_n w_i^{n-1}$ for $i = 1, \ldots, n-1$ then both sides will have the same minimizer for any choice of $p(x, y)$ and R-strict stability will hold. □

[2]Here we allow the relationship $w_i^n = 0, \forall i \neq n, w_n^n > 0$ that results for $\lambda_n = 0$.

As a corollary, we establish that all unweighted quasi-arithmetic means and quasi-medians are R-strictly stable (take $\lambda_n = 1$), which extends the results of [8, 19].

The next proposition shows that in the case of $p(x, y)$ being *differentiable*, this relationship between the weighting vectors is also the necessary condition for stability.

Proposition 3 *For a family $\{A_n\}_n$ of faithful penalty-based aggregation operators where the penalty function $p(x, y)$ is differentiable in y, R-strict stability holds if and only if the weighting convention described in Proposition 2 is satisfied.*

PROOF. From the previous proof we know that Eq. (8) must hold for stability. Given that $p(x, y)$ must be defined such that y takes a unique value in $[0, 1]$, for every fixed $\mathbf{x}$, we will have either $y = 0, y = 1$ or:

$$\frac{d}{dy}\left(\sum_{i=1}^{n-1} w_i^n p(x_i, y)\right) = \sum_{i=1}^{n-1} w_i^n \frac{d}{dy} p(x_i, y) = 0 \tag{9}$$

$$\frac{d}{dy}\left(\sum_{i=1}^{n-1} w_i^{n-1} p(x_i, y)\right) = \sum_{i=1}^{n-1} w_i^{n-1} \frac{d}{dy} p(x_i, y) = 0 \tag{10}$$

It follows from this that $\mathbf{w}^n$ and $\mathbf{w}^{n-1}$ can only differ by a scalar multiple and necessarily that

$$w_i^n = \lambda_n w_i^{n-1}, i = 1, \ldots, n-1.$$

□

This relationship between the weighting vectors had already been established by Gómez et al. for weighted arithmetic means, harmonic means, quadratic means and power means [16], however we see here that such a convention is necessary for the R-strict stability of all penalty-based functions with $p(x, y)$ differentiable. To show that differentiability of p is essential, we consider a case such as $p(x, y) = |x - y|$ where the weighting convention is still sufficient, but is not necessary to guarantee R-strict stability. Example 2 illustrates this last point.

Example 2 *Consider a lower weighted median (Definition 4) resulting from the penalty expression $p(x, y) = |x - y|$ and the 2-dimensional weighting vector, $\mathbf{w}^2 = (0.4, 0.6)$. We have the following two situations:*

$x_1 \geq x_2$, from which we obtain $\mathbf{u} = (0.4, 0.6)$ and $x_{(k)} = x_2$;
or
$x_2 \geq x_1$, which gives $\mathbf{u} = (0.6, 0.4)$ and $x_{(k)} = x_2$ again.
In fact, for any 2-dimensional weighting vector with $w_1^2 < w_2^2$, it will follow that $Med_{\mathbf{w}}(\mathbf{x}) = x_2$.

For 3 inputs, we then have $x_3 = x_2$ and any weighting vector with $w_1^3 < 0.5$ will result in a weighted median that is R-strictly stable with respect to the 2-dimensional case. For instance, the relationship between the weighting vectors $\mathbf{w}^2 = (0.4, 0.6)$ and $\mathbf{w}^3 = (0.45, 0.3, 0.25)$ is R-strictly stable for weighted medians, even though the ratio $w_1 : w_2$ is not preserved.

This merely shows that R-strict stability may not be the best indicator of consistency for penalty-based aggregation functions defined with respect to a non-differentiable penalty. When $p(x, y)$ is differentiable, however, and $w_n^n \neq 1$, R-strict stability is equivalent to the preservation of the ratios between each of the weights. We have the following useful corollaries that follow from Propositions 2 and 3.

Corollary 3 *All weighted quasi-arithmetic means are R-strictly stable if and only if for any sequence of weights $\mathbf{w}^n, \mathbf{w}^{n-1}$ it holds that*

$$w_i^n = (1 - w_n^n) w_i^{n-1}, i = 1, 2, \dots, n-1.$$

Proof. Direct from Proposition 3 with $\lambda_n = (1 - w_n^n)$ and the ability to model these functions in the form of Eq. (3). □

For quasi-weighted medians in the next proposition, we have only the one-directional *if* part.

Corollary 4 *All weighted quasi-medians are R-strictly stable if for any sequence of weights $\mathbf{w}^n, \mathbf{w}^{n-1}$ it holds that*

$$w_i^n = (1 - w_n^n) w_i^{n-1}, i = 1, 2, \dots, n-1.$$

Proof. Direct from Proposition 2 with $\lambda_n = (1 - w_n^n)$. □

Corollary 5 *A family of alpha-quantile operators is R-strictly stable with $\alpha = \frac{c}{1+c}$ provided c is fixed for all n.*

Proof. Direct from Proposition 2 with $p(x, y)$ defined as it is in Eq. (4) and equal weights for all i. □

The α-quantile operator includes special cases of the median (with $c = 1$) the maximum ($c = \infty$) and the minimum ($c = 0$), all of which can be defined with respect to its penalty expression with equal weights. This result does not extend to k-order statistics, however, since c depends on n and the penalty expressions would differ for A_n and A_{n-1}. We provide an example for the alpha-quantile operator.

Example 3 *Consider the α-quantile operator with $\alpha = \frac{3}{4}$. This has a penalty expression of*

$$p(x,y) = \begin{cases} x - y, & if\ x \leq y, \\ 3(y-x), & if\ x > y. \end{cases}$$

Suppose we have the input vector $\mathbf{x}_{i \neq n} = (0.2, 0.3, 0.7, 0.9)$. Any value in the interval $[0.7, 0.9]$ will minimize the overall penalty so we take the mid-point and have

$$A_{n-1}(0.2, 0.3, 0.7, 0.9) = 0.8.$$

For A_n we minimize the penalty with respect to the input vector,

$$\mathbf{x} = (0.2, 0.3, 0.7, 0.9, A_{n-1}(\mathbf{x}_{i \neq n})) = (0.2, 0.3, 0.7, 0.9, 0.8)$$

We then will have $A_n(\mathbf{x}) = 0.8$ since $p(0.8, 0.8) = 0$ and we already know that 0.8 minimizes the penalty of the original inputs in $\mathbf{x}$, so

$$A_n(0.2, 0.3, 0.7, 0.9, A_{n-1}(\mathbf{x}_{i \neq n})) = A_{n-1}(\mathbf{x}_{i \neq n})$$

as required.

The following example illustrates the application of the weighting convention in Corollary 3.

Example 4 *Consider the weighting vector $\mathbf{w}^4 = (\frac{3}{20}, \frac{5}{20}, \frac{8}{20}, \frac{4}{20})$. The R-strictly stable weighting triangle would require the relative ratios between each pair w_i^n, w_{i+1}^n to be preserved. It would be given by*

$$\begin{array}{ccccccc} & & & 1 & & & \\ & & \frac{3}{8} & & \frac{5}{8} & & \\ & \frac{3}{16} & & \frac{5}{16} & & \frac{8}{16} & \\ \frac{3}{20} & & \frac{5}{20} & & \frac{8}{20} & & \frac{4}{20} \\ & & & \cdots & & & \end{array}$$

Although Proposition 2 extends to a number of important aggregation functions including weighted means and quantile operators, it cannot be used to establish R-strict stability for more general aggregation functions that require a reordering step in their calculation such as the OWA function. We will now turn to penalty-based functions defined with respect to an auxiliary order-inducing variable.

4 R-strict stability for penalty-based aggregation functions with order inducing variables

In this section we will first present some general results, then investigate some specific weighting conventions associated with the OWA function. We also show how these considerations apply in the context of k-Nearest Neighbours (kNN) function approximation, where the output value of an unknown datum is predicted by averaging the closest k observed data. For the following considerations, we will focus on stability in the relationship between $\mathbf{w}^n$ and $\mathbf{w}^{n-1}$ as it will not always be possible to define stable weighting triangles.

4.1 General results

As stated in [16], the OWA function, in general is neither L- nor R-strictly stable. We have seen that some special cases, namely the α-quantile and weighted medians, which include the median, minimum and maximum, are R-strictly stable (and L-strictly stable since they are symmetric). We are interested in the conditions on weighting vectors associated with order-induced functions for R-strict stability.

Definition 10 *Given an order-inducing variable* $\mathbf{z}$ *where* z_i *is associated with the input* x_i*, a weighted order-induced aggregation function associates the weight* w_i *with* $x_{\sigma(i)}$ *where* $\sigma(\mathbf{x})$ *denotes a reordering of the inputs,* $\sigma(\mathbf{x}) = (x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{(n)})$ *such that* $z_{\sigma(1)} \geq z_{\sigma(2)} \geq \ldots \geq z_{\sigma(n)}$.

In the case of $z_i = x_i$, we will have inputs arranged in non-increasing order as they are for the standard OWA operator, while $z_i = i$ will lead to the x_i being associated with w_i as is the case for weighted means.

In general, it is possible to choose an order-inducing variable as some function of the x_i such that the relative ordering could change with the introduction of a new input. The following example illustrates this case.

Example 5 *We wish to define a consensus driven aggregation framework by allocating higher weight to values which are close to the mean. We therefore define the auxiliary order-inducing variable by*

$$z_i = 1 - \left| x_i - \frac{1}{n}\sum_{i=1}^{n} x_i \right| .$$

It is clear that the introduction of an extreme value could affect the relative ordering. Consider the inputs $\mathbf{x} = (0.1, 0.3, 0.8)$*. The mean is* 0.4 *so our input pairs for aggregation will be*

$$\langle \mathbf{x}, \mathbf{z} \rangle = (\langle 0.1, 0.7\rangle, \langle 0.3, 0.9\rangle, \langle 0.8, 0.6\rangle)$$

and we have $\sigma(\mathbf{x}) = (x_2, x_1, x_3)$*.*

We then introduce the input $x_4 = 0.8$*. The mean is now* 0.5 *and we have*

$$\langle \mathbf{x}, \mathbf{z} \rangle = (\langle 0.1, 0.6\rangle, \langle 0.3, 0.8\rangle, \langle 0.8, 0.7\rangle, \langle 0.8, 0.7\rangle).$$

The order induced by $\mathbf{z}$ *is now* $\sigma(\mathbf{x}) = (x_2, x_4, x_3, x_1)$ *and the relative ordering of* x_1 *and* x_3 *has changed.*

In order to avoid such situations, we will limit ourselves to order-inducing variables that satisfy the following definition of *order consistency* based on the definition given in [1].

Definition 11 *Consider two pairs of vectors* $\langle \mathbf{x}, \mathbf{z} \rangle$ *and* $\langle \mathbf{x}', \mathbf{z}' \rangle$ *where* $\mathbf{z}'$ *is obtained from* z *by considering an additional input* $\langle x_{n+1}, z_{n+1} \rangle$*. An inducing variable is order consistent if whenever it holds that* $z_i \geq z_j$*, then it also holds that* $z'_i \geq z'_j$*.*

In other words, a new pair $\langle x_{n+1}, z_{n+1} \rangle$ does not affect the relative ordering of $\langle \mathbf{x}, \mathbf{z} \rangle$.

For penalty-based aggregation functions, we consider the inputs $\langle x_i, z_i \rangle$ such that the weight w_i is associated with the $\sigma(i)$-th penalty $p(x_{\sigma(i)}, y)$.

From the n inputs we obtain a reordering according to $\mathbf{z}$ and let $x_{\sigma(j)} = x_n$. This gives the following reordered input vectors for n and $n-1$ arguments respectively,

$$\sigma(\mathbf{x}) = (x_{\sigma(1)}, \ldots, x_{\sigma(j-1)}, x_n, x_{\sigma(j+1)}, \ldots, x_{\sigma(n)})$$

$$\sigma(\mathbf{x}_{i\neq n}) = (x_{\sigma(1)}, \ldots, x_{\sigma(j-1)}, x_{\sigma(j+1)}, \ldots, x_{\sigma(n)})$$

The penalty expressions replacing Eq. (5) and Eq. (6) will then be given by

$$A_n(\mathbf{x}) = \arg\min_y \left(\sum_{i=1}^{j-1} w_i^n p(x_{\sigma(i)}, y) + w_j^n p(x_n, y) + \sum_{i=j+1}^{n} w_i^n p(x_{\sigma(i)}, y) \right) \tag{11}$$

$$A_{n-1}(\mathbf{x}_{i\neq n}) = \arg\min_y \left(\sum_{i=1}^{j-1} w_i^{n-1} p(x_{\sigma(i)}, y) + \sum_{i=j+1}^{n} w_{i-1}^{n-1} p(x_{\sigma(i)}, y) \right) \tag{12}$$

Note that the effect of x_n being ordered in the j-th position rather than the n-th is that the weighting indices in A_{n-1} are shifted by 1 for all $i > j$ (see Corollary 1), e.g. if x_n was reordered to $x_{\sigma(3)}$ then the weight w_5^{n-1} would be associated with $x_{\sigma(6)}$ in Eq. (12) while w_5^n would be associated with $x_{\sigma(5)}$ in Eq. (11). For ease in interpreting the relationship between the two equations, let us express $\mathbf{w}^{n-1}$ in terms of an $(n-1)$-dimensional weighting vector $\mathbf{u} = (u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_n)$ where

$$u_i = \begin{cases} \lambda_n w_i^{n-1}, i < j, \\ \lambda_n w_{i-1}^{n-1}, i > j. \end{cases}$$

Eqs. (11) and (12) then can be written

$$A_n(\mathbf{x}) = \arg\min_y \left(\sum_{i=1, i\neq j}^{n} w_i^n p(x_{\sigma(i)}, y) + w_j^n p(x_n, y) \right) \tag{13}$$

$$A_{n-1}(\mathbf{x}_{i\neq n}) = \arg\min_y \left(\sum_{i=1, i\neq j}^{n} u_i p(x_{\sigma(i)}, y) \right) \tag{14}$$

We see that where we know the value of j (the position in which x_n is aggregated), we will have an analogous situation to standard orderings and can draw upon the result from Corollary 1. Although in the previous analysis we were able to remove the term associated with $x_n = A_{n-1}(\mathbf{x}_{i\neq n})$, here it will not be the case in general that

$x_n = x_{\sigma(j)}$ for *all* $\mathbf{x}$, i.e. the value of j may be dependent on the aggregated value $A_{n-1}(\mathbf{x}_{i \neq n})$. Example 6 helps to illustrate this last point.

Example 6 *Consider an OWA operator with* $\mathbf{w} = (0.5, 0.3, 0.2)$. *The aggregated value* $A_{n-1}(0.8, 0.3, 0.1) = 0.51$ *so when we include 0.51 and aggregate* $A_n(0.8, 0.3, 0.1, 0.51)$, *we have* $\sigma(\mathbf{x}) = (0.8, 0.51, 0.3, 0.1)$ *and* $x_n = x_{\sigma(2)}$. *On the other hand, aggregating* $A_{n-1}(0.8, 0.7, 0.1) = 0.63$ *means we would then aggregate* $A_n(0.8, 0.7, 0.1, 0.63)$ *and* $\sigma(\mathbf{x}) = (0.8, 0.7, 0.63, 0.1)$, *i.e.* $x_n = x_{\sigma(3)}$. *So for this function, the value of* j *when we determine* $x_n = x_{\sigma(j)}$ *will depend on the input values.*

However, there are situations where the value of j may be the same for all input vectors, e.g. in the case of k-order statistics.

We begin with the following proposition, which states that a family of order-induced aggregation operators will be R-strictly stable if we can define weighting vectors consistently according to j.

Proposition 4 *Given an order consistent inducing variable* $\mathbf{z}$ *and a family of penalty-based aggregation operators with*

$$P(\mathbf{x}, y) = \sum_{i=1}^{n} w_i p(x_{\sigma(i)}, y),$$

if the weighting vectors $\mathbf{w}^n$ *and* $\mathbf{w}^{n-1}$ *associated with the penalty expressions of* A_n *and* A_{n-1} *satisfy:*

$$w_i^n = \begin{cases} \lambda_n w_i^{n-1}, i < j, \\ \lambda_n w_{i-1}^{n-1}, i > j, \end{cases}$$

with $\lambda_n \geq 0$ *a constant for all possible* j *such that* $z_n = z_{\sigma(j)}$, *then the function is considered* R*-strictly stable.*

Proof. We can establish this analogously to Proposition 2 from the similarity between Eqs. (5), (6) and Eqs. (13), (14). Since $z_n = z_{\sigma(j)}$, it will follow that $x_n = A_{n-1}(\mathbf{x}_{i \neq n})$ will be associated with the weight w_j^n and $p(x_{\sigma(j)}, y) = 0$, we then have the same requirement as Eq. (8) with a shift in the indexing for w_i^{n-1} for $i > j$. □

To show how this proposition applies, we will extend the previous example.

Example 7 *(Cont'd from Example 6) We wish to extend the OWA with* $\mathbf{w}^{n-1} = (0.5, 0.3, 0.2)$ *to include a 4th input. In order to define* $\mathbf{w}^n$ *such that the relationship between* $\mathbf{w}^{n-1}$ *and* $\mathbf{w}^n$ *is R-strictly stable, it should hold that*

$$OWA(0.8, 0.3, 0.1, 0.51) = 0.51 \text{ and } OWA(0.8, 0.7, 0.1, 0.63) = 0.63$$

(and the same for any $OWA(\mathbf{x}_{i\neq 4})$ *used as the input* x_4*). In the first case, where the aggregated value leads to* $j = 2$*, we require*

$$\begin{aligned} w_1^n &= \lambda_n 0.5, \\ w_3^n &= \lambda_n 0.3, \\ w_4^n &= \lambda_n 0.2, \end{aligned} \tag{15}$$

while for the latter case when $j = 3$ *we require*

$$\begin{aligned} w_1^n &= \lambda_n 0.5, \\ w_2^n &= \lambda_n 0.3, \\ w_4^n &= \lambda_n 0.2. \end{aligned} \tag{16}$$

We hence note from Eq. (15) and Eq. (16) that w_2^n *and* w_3^n *should be equal. This leads to the weighting vector* $\mathbf{w} = (\frac{5}{13}, \frac{3}{13}, \frac{3}{13}, \frac{2}{13})$ *and* $\lambda_n = 1 - \frac{3}{13}$*. Given that the output of this OWA is bounded such that* j *will only ever be equal to 2 or 3, R-strict stability will hold for all 3-dimensional input sets.*

It will not always be possible to define weighting vectors in such a way for order-induced functions. For instance, an OWA with $\mathbf{w} = (0.4, 0.3, 0.2, 0.1)$ has no corresponding R-strictly stable 5-dimensional weighting vector. The output could lead to x_n being aggregated in any position $j = 2, 3, 4$, so the reasoning from Example 7 would lead to the requirement of $w_2^n = w_3^n = w_4^n$ which is not possible since $\lambda_n 0.3 \neq \lambda_n 0.2$.

We do, on the other hand, have the following corollary.

Corollary 6 *A family of order-induced aggregation functions can be defined such that R-strict stability holds if* $A_{n-1}(\mathbf{x}_{i\neq n}) = x_{\sigma(j)}$ *with* j *fixed for any given* n*.*

Proof. Direct from Proposition 4. □

From this we establish that all k-order statistics are R-strictly stable since $j = k$ for all n, i.e. $A_{n-1}(\mathbf{x}_{i\neq n}) = x_{\sigma(k)}$. We hence have

$\lambda_n = 0$ for all n and $w_j^n = 1$. Proposition 4 can also be applied to families of weighted means where each additional input is not necessarily aggregated in the n-th position. For instance, suppose we have $\mathbf{w}^3 = (0.7, 0.1, 0.2)$ and we want to define a weighting vector for the case when x_2 is missing. The relationship between $\mathbf{w}^{n-1} = (\frac{7}{9}, \frac{2}{9})$ and $\mathbf{w}^n$ is not considered R-strictly stable according to Proposition 2, however if the order rearrangement is taken into account using Proposition 4, we can consider it to be R-strictly stable with $\mathbf{z}$ defined such that $z_3 = z_{\sigma(2)}$. This provides us with a framework which allows us to extend the results of R-strict stability to the notion of stability with respect to the j-th input and an ordering variable.

In general, if we cannot predict the position j to which $x_n = A_{n-1}(\mathbf{x}_{i\neq n})$ is ordered, R-strict stability will require that the relationship between the weighting vectors specified in Proposition 4 holds with j taking multiple values as it did in Example 7. This gives us the following family of weighting vectors.

Corollary 7 *Consider an order-consistent induced aggregation function and the pair $\langle x_n, z_n\rangle$ such that $x_n = A_{n-1}(\langle \mathbf{x}_{i\neq n}, \mathbf{z}_{i\neq n}\rangle)$ and z_n is bounded relative to $\sigma(\mathbf{z})$, i.e.*

$$z_{\sigma(a)} \geq z_n \geq z_{\sigma(b)}.$$

If it holds that $w_{a+1}^{n-1} = w_{a+2}^{n-1} = \ldots = w_{b-2}^{n-1}$, then $\mathbf{w}^n$ can be defined according to Proposition 4 such that the function is R-strict stable.

When the penalty $p(x_n, y)$ is associated with the weight w_j^n, the previous indexation of weights for $i > j$ shifts across by 1. In order for the proportional weighting allocated to the inputs to be maintained for multiple j, a necessary requirement is that the weights distributed from $j = a+1$ to $j = b-1$ are the same, so for $\mathbf{w}^n$, we will have $w_{a+1}^n = w_{a+2}^n = \ldots = w_{b-1}^n$.

Proof. We assume that there exists an input vector $\mathbf{x}_{i\neq n}$ such that $z_n = z_{\sigma(j)}$ and that the weights satisfy Proposition 4. In particular, note that $w_{j-1}^n = \lambda_n w_{j-1}^{n-1}$. If there exists a set of inputs $\mathbf{x}'_{i\neq n}$ such that $z'_n = z'_{\sigma(j-1)}$, then we require

$$w_i^n = \begin{cases} \lambda_n w_i^{n-1}, & i < j-1, \\ \lambda_n w_{i-1}^{n-1}, & i > j-1, \end{cases}$$

with $\lambda \geq 0$ as a constant. In this case we have $w_j^n = \lambda_n w_{j-1}^{n-1}$ and we see that for both cases to simultaneously satisfy the weighting convention,

$$w_j^n = w_{j-1}^n = \lambda_n w_{j-1}^{n-1}.$$

By extending this reasoning, we determine that the weighting relationship will need to hold for all possible j such that $A_{n-1}(\mathbf{x}_{i\neq n}) = x_{\sigma(j)}$. □

This means weighting conventions such as that used for the olympic average, $\mathbf{w} = (0, w, w, \dots, w, 0), w = \frac{1}{n-2}$ and in fact any weighting vector $\mathbf{w} = (0, \dots, 0, w_a, w, w, \dots, w, w_b, 0, \dots, 0)$ (including the *trimmed means*) can be considered R-strictly stable since the output will always be bounded between the inputs associated with w_a and w_b. A pair of weighting vectors that satisfy R-strict stability according to Corollary 7 are provided in Example 8.

Example 8 *Consider an OWA function with 6- and 7-dimensional weighting vectors defined as follows,*

$$\mathbf{w}^6 = (0, \frac{2}{6}, \frac{1}{6}, \frac{1}{6}, \frac{2}{6}, 0).$$

$$\mathbf{w}^7 = (0, \frac{2}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{2}{7}, 0)$$

We note that Proposition 7 is satisfied for $j = 3, 4, 5$. Let us assume that the inputs have been presorted into non-increasing order. We know that $OWA(x_1, x_2, \dots, x_6) = x_7$ will be bounded between x_2 and x_5 so, when checking for R-strict stability, in the calculation of $OWA(x_1, \dots, x_6, x_7)$ we will have

$$OWA(\mathbf{x}) = \frac{2}{7}x_2 + \frac{1}{7}x_3 + \frac{1}{7}x_4 + \frac{2}{7}x_5 + \frac{1}{7}x_7$$

regardless of whether $x_7 = OWA(\mathbf{x}_{i\neq 7})$ is aggregated in position $j = 3, 4, 5$ according to its value.

We will now turn our attention to special ways of defining OWA weights and the conditions for R-strict stability.

4.2 Defining weights for OWA operators

In the case of OWA operators, $z_i = x_i$ and we will simply use the notation $x_{(i)}$ to denote the i-th largest input. We can determine from Corollary 7 that weighting vectors are R-strictly stable if

$$x_{(a)} > OWA(\mathbf{x}_{i\neq n}) > x_{(b)}$$

holds for some a, b and $w_i = w_j, \forall i, j \in \{a+1, \ldots, b-1\}$. This includes k-order statistics with $a + 1 = b - 1 = k$.

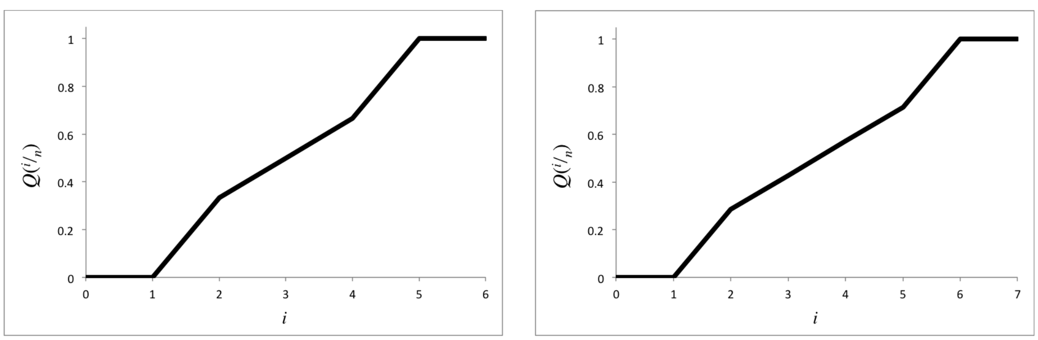
In [21,22], Yager proposed the use of Basic Unit-interval Monotone (BUM) functions $Q : [0, 1] \to [0, 1], Q(0) = 0, Q(1) = 1$, or Regular Increasing Monotone (RIM) quantifiers in order to define the weighting vectors for OWA functions. For a given Q and n, the weights are calculated using:

$$w_i^n = Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right).$$

Examples of RIM quantifiers include "for all" with $Q(1) = 1, 0$ otherwise; "there exists" with $Q(0) = 0, 1$ otherwise; and other concepts using

$$Q_{a,b}(t) = \begin{cases} 0, & \text{if } t \leq \frac{a}{n}, \\ \frac{nt-a}{b-a}, & \text{if } \frac{a}{n} < t < \frac{b}{n}, \\ 1, & \text{if } t \geq \frac{b}{n}. \end{cases}$$

In particular, OWA functions can model various linguistic quantifiers, e.g. suppose $n = 5$, "most" can be modeled with the vector $\mathbf{w} = (1/3, 1/3, 1/3, 0, 0)$ and "80% of" with $\mathbf{w} = (0, 0, 0, 1, 0)$.



(a) (b)

Figure 1: RIM quantifiers corresponding with the weighting vectors of 6- and 7-dimensions from Example 8.

For quantifier based aggregation, R-strict stability requires the successive weights to be equal, $w_i = w_j$, i.e. that

$$Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right) = Q\left(\frac{j}{n}\right) - Q\left(\frac{j-1}{n}\right),$$

for $i, j \in \{a+1, \ldots, b-1\}$. In other words, we require that $Q(t)$ be linearly increasing over the interval $[\frac{a}{n}, \frac{b-1}{n}]$. This means the $Q_{a,b}(t)$ operators are a good candidate for forming the basis of R-strictly stable aggregation, however, it should be kept in mind how the quantifier is defined as n changes. Figure 1 shows the RIM quantifiers from Example 8. For the 7-dimensional vector, the function has a constant gradient between $i = 2$ and $i = 5$ since x_7 could be aggregated in any position from $j = 3$ to $j = 5$.

4.3 Application to kNN

In k-Nearest Neighbors (kNN) function approximation, we predict the value y_0 for an unknown input vector $\mathbf{x}_0$ by averaging the y-values of the closest k observed data. We hence aggregate the inputs $\langle y_i, z_i \rangle$ where y_i denotes the observed function value for a given $\mathbf{x}_i$ and z_i is an auxiliary variable indicating the proximity, e.g using inverse Euclidean distance with $z_i = \frac{1}{||\mathbf{x}_i - \mathbf{x}_0||}$. It is clear that $\mathbf{z}$ is order-consistent, since the introduction of a new datum $\langle y_i, z_i \rangle$ will not affect the relative distances between $\mathbf{x}_0$ and each of the $\mathbf{x}_i$. We consider the problem of defining R-strictly stable weights for the set of k inputs used in kNN. In particular, we are interested in the stability of the weights when the size of k is incremented by 1, i.e. we introduce the datum $(\mathbf{x}_{k+1}, y_{k+1})$.

In standard kNN, we take the arithmetic mean of the nearest k data. Clearly this corresponds with using equal weights and the aggregation is R-strictly stable.

One weighted extension of kNN weights the neighbors according to inverse distance, i.e.

$$w_i = \frac{\left(\frac{1}{||\mathbf{x}_{(i)} - \mathbf{x}_0||}\right)}{\sum\limits_{i=1}^{k} \frac{1}{||\mathbf{x}_{(i)} - \mathbf{x}_0||}}. \tag{17}$$

Now we consider the inclusion of a new datum as we increase the value of k. In this case, $\mathbf{x}_{k+1}$ will be further away from $\mathbf{x}_0$ than the data originally included in the aggregation, and hence the auxiliary variable z_{k+1} denoting proximity will be the $(k+1)$-th largest, i.e. $z_{k+1} = z_{(k+1)}$.

We can show that R-strict stability is satisfied automatically when we use Eq. (17) to define the weights. The calculation of each w_i will result in the denominator sum increasing by $\frac{1}{||\mathbf{x}_{k+1}-\mathbf{x}_0||}$, which means that we have

$$\lambda_{k+1} = \frac{\sum_{i=1}^{k} \frac{1}{||\mathbf{x}_{(i)}-\mathbf{x}_0||}}{\sum_{i=1}^{k+1} \frac{1}{||\mathbf{x}_{(i)}-\mathbf{x}_0||}},$$

and the new weights will satisfy the conditions in Proposition 4.

We provide a numerical example before concluding the paper.

Example 9 *Suppose we are trying to predict the output value* y_0 *for a function* f *when taking the inputs* $\mathbf{x}_0 = (0.2, 0.5, 0.85, 0.92)$*. The input values and output for the 4 closest observations are given in Table 1 with their inverse Euclidean distances* $z_i = \frac{1}{||\mathbf{x}_i-\mathbf{x}_0||}$ *given in the last column.*

For $k = 3$*, the closest 3 data points are* $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ *and the weighting vector we obtain from Eq. (17) is* $\mathbf{w}^3 = (0.462, 0.273, 0.265)$*. The predicted output value would then be (to 2 decimal places)*

$$f(\mathbf{x}_0) = 0.462 \times 0.75 + 0.273 \times 0.73 + 0.265 \times 0.49 = 0.68.$$

If we then increase k *to 4, we will include* $\mathbf{x}_4$ *and need to extend the weighting vector. Eq. (17) gives us* $\mathbf{w}^4 = (0.368, 0.217, 0.211, 0.204)$*. Note that* w_i^4 *for* $i = 1, 2, 3$ *can each be obtained by multiplying the respective* w_i^3 *by* $(1-w_4^4)$*. Our predicted output for* $k = 4$ *(to 2 decimal places is)*

$$f(\mathbf{x}_0) = 0.368 \times 0.75 + 0.217 \times 0.73 + 0.211 \times 0.49 + 0.204 \times 0.68 = 0.68,$$

and we see that R-strict stability is satisfied since the new input was equal to the aggregated value for when $k = 3$ *and* $f(\mathbf{x}_0)$ *remained unchanged.*

$\mathbf{x}_i$	x_{i1}	x_{i2}	x_{i3}	x_{i4}	y_i	z_i
$\mathbf{x}_1$	0.21	0.44	0.87	1.00	0.75	10.621
$\mathbf{x}_2$	0.30	0.47	0.93	1.00	0.73	6.271
$\mathbf{x}_3$	0.30	0.52	0.76	0.84	0.49	6.096
$\mathbf{x}_4$	0.16	0.46	0.87	0.76	0.68	5.907

Table 1: Input vectors for Example 9 with their output y_i and inverse Euclidean distance z_i to the unknown data point $\mathbf{x}_0 = (0.2, 0.5, 0.85, 0.92)$.

5 Conclusion

We adopted definitions of stability of families of aggregation functions with respect to input cardinality from [19], and extended their results to quasi-arithmetic means, quasi-medians and other penalty-based aggregation functions.

We have looked thoroughly at conditions on weights for a number of important aggregation functions. We established that relations between the weighting vectors from [19] are sufficient for all weighted faithful penalty-based functions, but are necessary only for functions with differentiable penalties, thus leaving the case of weighted medians apart.

In particular, we note that:

- Weighted means defined with respect to a weighted penalty expression where the penalty is differentiable will be R-strictly stable if and only if the ratio between w_i^n and w_i^{n-1} is $(1-w_n^n)$. This result can also extend to j-th position stability with an appropriate re-indexing of the weights. It is hence possible to define R-strictly stable weighting triangles for all quasi-arithmetic means;
- The same weighting convention is sufficient to ensure stability for all weighted penalty-based aggregation functions. This means we can define R-strictly stable weighting triangles for all weighted quasi-medians and α-quantile operators.

We have also considered the notion of R-strict stability for weighted aggregation functions that are based on a reordering of the inputs. This led us to defining R-strictly stable weighting vectors of OWA and induced OWA operators. In particular, we concluded that:

- If we know the relative position of the aggregated value with respect to the inputs and ordering method, it is possible to define R-strictly stable weighting triangles. All k-order statistics, including the special cases of the maximum, minimum are R-strictly stable;
- If the relative position of the aggregated value with respect the inputs can change but is *bounded*, the R-strictly stable weighting vectors are required to have equal weights for all positions that the aggregated value could take. OWA functions such as the olympic average and trimmed means are hence R-strictly stable, since the aggregated input will always be weighted the same

when appended to the original input vector. This result can also be applied to OWAs with weights defined by ab-quantifiers.

An interesting application to the kNN machine learning method was also illustrated.

References

[1] A. Amo, J. Montero, and E. Molina. Representation of consistent recursive rules. *European Journal of Operational Research*, 130:29–53, 2001.

[2] G. Beliakov, T. Calvo, and S. James. On Lipschitz properties of generated aggregation functions. *Fuzzy Sets and Systems*, 161(10):1437–1447, 2010.

[3] G. Beliakov, S. James, J. Mordelová, T. Rückschlossová, and R. R. Yager. Generalized Bonferroni mean operators in multi-criteria aggregation. *Fuzzy Sets and Systems*, 161:2227–2242, 2010.

[4] G. Beliakov, A. Pradera, and T. Calvo. *Aggregation Functions: A Guide for Practitioners.* Springer, Heidelberg, 2007.

[5] H. Bustince, E. Barrenechea, T. Calvo, S. James, and G. Beliakov. Consensus in multi-expert decision making problems using penalty functions defined over a Cartesian product of lattices. *Information Fusion*, doi:10.1016/j.inffus.2011.10.002, 2011.

[6] T. Calvo. Two ways of generating extended aggregation functions. In *Proc. IPMU'98, Paris, France*, pages 825–831. 1998.

[7] T. Calvo and G. Beliakov. Aggregation functions based on penalties. *Fuzzy Sets and Systems*, 161:1420–1436, 2010.

[8] T. Calvo, M. Carbonell, and P. Canet. Medias casilineales ponderadas extendidas. In *Proc. Estylf'97, Tarragona, Spain*, pages 33–38. 1997.

[9] T. Calvo and G. Mayor. Remarks on two types of extended aggregation functions. *Tatra Mount. Math. Publ.*, 16(1):235–255, 1999.

[10] T. Calvo, G. Mayor, J. Torrens, J. Suñer, M. Mas, and M. Carbonell. Generation of weighting triangles associated with aggregation functions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 8(4):417–451, 2000.

[11] T. Calvo and R. Mesiar. Stability of aggregation operators. In *2nd Conf. of the Europ. Soc. for Fuzzy Logic and Technology*, pages 475–478, Leicester, 2001.

[12] T. Calvo, R. Mesiar, and R. R. Yager. Quantitative weights and aggregation. *IEEE Trans. on Fuzzy Systems*, 12:62–69, 2004.

[13] V. Cutello and J. Montero. Recursive families of OWA operators. In *Proc. of FUZZIEEE*, pages 1137–1141, Piscataway, NJ, 1994.

[14] M. Gągolewski and P. Grzegorzewski. Arity-monotonic extended aggregation operators. In E. Hüllermeier, R. Kruse, and F. Hoffmann, editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems (Communications in Computer and Information Sciences 80)*. Springer-Verlag, Heidelberg.

[15] M. Gągolewski and P. Grzegorzewski. Possibilistic analysis of arity-monotonic aggregation operators and its relation to bibliometric impact assessment of individuals. *International Journal of Approximate Reasoning*, 52:131 –1324, 2011.

[16] D. Gómez, J. Montero, J. Tinguaro Rodríguez, and K. Rojas. Stability in aggregation operators. In *Proc. of IPMU*, Catania, Italy, 2012.

[17] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap. *Aggregation Functions.* Cambridge University Press, Cambridge, 2009.

[18] A. Pradera, E. Trillas, and T. Calvo. A general class of triangular norm-based aggregation operators: quasi-linear T-S operators. *Intl. J. of Approximate Reasoning*, 30:57–72, 2002.

[19] K. Rojas, D. Gómez, J. Tinguaro Rodríguez, and J. Montero. Some properties of consistency in the families of aggregation operators. In B. De Baets et al., editor, *Eurofuse 2011, AISC 107*, pages 169–176. Springer, 2011.

[20] Y. Torra and V. Narukawa. *Modeling Decisions. Information Fusion and Aggregation Operators.* Springer, 2007.

[21] R. R. Yager. Connectives and quantifiers in fuzzy sets. *Fuzzy Sets and Systems*, 40:39–76, 1991.

[22] R. R. Yager. Families of OWA operators. *Fuzzy Sets and Systems*, 59:125148, 1993.

[23] R. R. Yager and D. P. Filev. Induced ordered weighted averaging operators. *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, 20:141–150, 1999.

[24] R. R. Yager and A. Rybalov. Noncommutative self-identity aggregation. *Fuzzy Sets and Systems*, 85:73–82, 1997.

[25] H.-J. Zimmermann and P. Zysno. Latent connectives in human decision making. *Fuzzy Sets and Systems*, 4:37–51, 1980.